NOISE LEVEL ESTIMATION USING WEAK TEXTURED PATCHES OF A SINGLE NOISY IMAGE

Xinhao Liu   Masayuki Tanaka   Masatoshi Okutomi

Department of Mechanical and Control Engineering, Tokyo Institute of Technology

ABSTRACT

A patch-based noise level estimation algorithm is proposed in this paper, with patches generated from a single noisy image. One can easily estimate the noise level from image patches using principal component analysis (PCA) if the image comprises only weak textured patches. The challenge for patch-based noise level estimation is how to select weak textured patches from a noisy image. As described in this paper, we propose a novel algorithm to select weak textured patches from a single noisy image based on the gradients of the patches and their statistics. Then we estimate the noise level from the selected weak textured patches using PCA. We demonstrate experimentally that the proposed noise level estimation algorithm outperforms the state-of-the-art algorithm.

***Index Terms*—** noise level estimation, weak textured patch, image gradient, PCA

1. INTRODUCTION AND RELATED WORK

A noise level or standard deviation of the Gaussian noise is necessary for many image-processing algorithms and applications. For instance, many denoising algorithms operate on the assumption that the noise level is known a priori, which in fact is not valid in practical circumstances. Image segmentation and smoothing can be improved significantly if the noise level is known. However, in current studies, the noise level is usually provided manually. It is therefore a challenge to assign an accurate noise level for a variety of input images, especially for those with rich textures.

Many algorithms [1, 2, 3, 4] have been proposed for gray-level image noise level estimation. Generally they are classifiable into patch-based and filter-based approaches, or some combination of them. Tai *et al.* [3] proposed a filter-based noise estimation method that uses a Laplacian operator to suppress the image structure and adaptive edge detection to exclude pixels associated with edges. The main difficulty inherent in filter-based methods is that the difference between the original and filtered image is assumed to be the noise, but this assumption is not true for images with complex structures or details. Shin *et al.* [4] proposed a patch-based algorithm in which an image is split into numerous patches. Then smooth patches are selected. Subsequently, the noise level is computed from the selected patches. The main issue of patch-based methods is how to identify the weak textured or smooth patches for various scenes in the presence of Gaussian noise. Different from the methods described above, Zoran and Weiss [5] reported that the changes in kurtosis values of noisy images are due to the noise presented in the image and the change can be modeled to estimate the noise level. Their algorithm can handle the low noise level situation quite well and achieves the state-of-the-art result.

Existing texture strength measures such as variance can not reflect the underlying texture strength correctly in the presence of the noise. In this paper, we propose a patch-based noise estimation algorithm using principal component analysis (PCA) and a novel texture strength metric to select the weak textured patches. We investigate the relation between the proposed metric and the noise level σ_n. Then we propose an iterative framework to select weak textured patches from the noisy image with different noise levels. The experimental results demonstrate that our estimation method works well for various scenes and that it outperforms the current leading methods.

2. PROPOSED ALGORITHM

2.1. Noise Level Estimation Based on PCA

After decomposing the image into overlapping patches, we can write the image model as

$$\mathbf{y}_i = \mathbf{z}_i + \mathbf{n}_i \,, \tag{1}$$

where $\mathbf{z}_i$ is the original image patch with the i-th pixel at its center written in a vectorized format and $\mathbf{y}_i$ is the observed vectorized patch corrupted by i.i.d zero-mean Gaussian noise vector $\mathbf{n}_i$. The goal of noise level estimation is to calculate the unknown standard deviation σ_n given only the observed noisy image.

The image patches can be regarded as data in Euclidean space. Let us consider the variance of the data projected onto a certain axis. We can define the direction of the axis using the unit vector $\mathbf{u}$. Assuming that the signal and the noise are uncorrelated, the variance of the projected data on direction $\mathbf{u}$ can be expressed as:

$$V(\mathbf{u}^T\mathbf{y}_i) = V(\mathbf{u}^T\mathbf{z}_i) + \sigma_n^2 \,, \tag{2}$$

where $V(\mathbf{y}_i)$ represents the variance of the dataset $\{\mathbf{y}_i\}$, σ_n is the standard deviation of the Gaussian noise. We define the minimum variance direction $\mathbf{u}_{\min}$ as

$$\mathbf{u}_{\min} = \arg\min_{\mathbf{u}} V(\mathbf{u}^T\mathbf{z}_i) = \arg\min_{\mathbf{u}} V(\mathbf{u}^T\mathbf{y}_i) \,. \tag{3}$$

Following the same manner of the maximum variance formulation in [6], the minimum variance direction is calculable using the PCA. The minimum variance direction is the eigenvector associated to the minimum eigenvalue of the covariance matrix defined as

$$\mathbf{\Sigma_y} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{y}_i - \boldsymbol{\mu})(\mathbf{y}_i - \boldsymbol{\mu})^T \,, \tag{4}$$

where N is the data number and $\boldsymbol{\mu}$ is the average of the dataset $\{\mathbf{y}_i\}$.

The variance of the data projected onto the minimum variance direction equals the minimum eigenvalue of the covariance matrix. Then we can derive the equation

$$\lambda_{min}(\mathbf{\Sigma_y}) = \lambda_{min}(\mathbf{\Sigma_z}) + \sigma_n^2 \,, \tag{5}$$

(a) $s_1^2 = 21830.30$ (b) $s_1^2 = 14109.05$ (c) $s_1^2 = 31.12$

Fig. 1. Maximum eigenvalue s_1^2 of patches of different types, weak textured patches have smaller values.

where $\mathbf{\Sigma_y}$ signifies the covariance matrix of the noisy patch $\mathbf{y}_i$, $\mathbf{\Sigma_z}$ denotes the covariance matrix of the noise-free patch $\mathbf{z}_i$, and $\lambda_{\min}(\mathbf{\Sigma})$ represents the minimum eigenvalue of the matrix $\mathbf{\Sigma}$.

The noise level can be estimated easily if we can decompose the minimum eigenvalue of the covariance matrix of the noisy patches as Eq.(5). However, the decomposition problem is an ill-posed problem because the minimum eigenvalue of the covariance matrix of the noise-free patches is unknown. Although this decomposition problem is an ill-posed problem in general, we can estimate the noise level if we can select weak textured patches from the noisy images as described below.

The weak textured patches are known to span only low-dimensional subspace. The minimum eigenvalue of the covariance matrix of such weak textured patches is approximately zero. Then, the noise level can be estimated simply as

$$\hat{\sigma}_n^2 = \lambda_{\min}(\mathbf{\Sigma_{y'}})\,, \tag{6}$$

where $\mathbf{\Sigma_{y'}}$ is the covariance matrix of the selected weak textured patches.

Consequently, we can estimate the noise level easily if we can select the weak textured patches from the noisy image. However, the weak textured image patch selection also presents a challenging problem, as discussed in the next subsection.

2.2. Weak Textured Patch Selection

Zhu and Milanfar [7] demonstrated that the image structure can be measured effectively by the gradient covariance matrix. The gradient covariance matrix, $\mathbf{C_y}$, for the image patch $\mathbf{y}$ is defined as:

$$\begin{aligned}\mathbf{C_y} &= \mathbf{G_y}^T\mathbf{G_y}\\ \mathbf{G_y} &= \begin{bmatrix}\mathbf{D}_h\mathbf{y} & \mathbf{D}_v\mathbf{y}\end{bmatrix}\,,\end{aligned} \tag{7}$$

where $\mathbf{D}_h$ and $\mathbf{D}_v$ respectively represent the matrices to represent horizontal and vertical derivative operators. Much information about the image patch can be reflected by the eigenvalue and eigenvector of the gradient covariance matrix.

$$\mathbf{C_y} = \mathbf{V}\begin{bmatrix}s_1^2 & 0\\ 0 & s_2^2\end{bmatrix}\mathbf{V}^T. \tag{8}$$

The maximum eigenvalue of the gradient covariance matrix s_1^2 reflects the strength of the dominant direction of that patch. The larger maximum eigenvalue reflects the richer texture. For this study, we use this maximum eigenvalue of the gradient covariance matrix as the quantitative measure for the texture strength of the image patches. Fig.1 shows three patches with different maximum eigenvalues of the gradient covariance matrix. It might be readily apparent that a smaller maximum eigenvalue of the gradient covariance matrix indicates the smoother or the weaker textured patch.

Unfortunately, the eigenvalue of the gradient covariance matrix is sensitive to noise. However, we must select the weak textured patches from the noisy image. Therefore, we analyze how the Gaussian noise affects the eigenvalue of the gradient covariance matrix. Considering a perfectly flat patch with N pixels where the Gaussian noise with standard deviation σ_n is added, the noisy flat patch $\mathbf{y}$ can be represented as

$$\mathbf{y} = \mathbf{f} + \mathbf{n}\,, \tag{9}$$

where $\mathbf{f}$ and $\mathbf{n}$ respectively represent the perfectly flat patch and Gaussian noise. Because the gradients of the perfectly flat patch are zero, following the derivation in [7] we can calculate the expected gradient covariance matrix of the noisy flat patch as

$$\begin{aligned}E(\mathbf{C_y}) = E(\mathbf{C_n}) &= E(\begin{bmatrix}\mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_h\mathbf{n} & \mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_v\mathbf{n}\\ \mathbf{n}^T\mathbf{D}_v^T\mathbf{D}_h\mathbf{n} & \mathbf{n}^T\mathbf{D}_v^T\mathbf{D}_v\mathbf{n}\end{bmatrix})\\ &= \begin{bmatrix}E(\mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_h\mathbf{n}) & 0\\ 0 & E(\mathbf{n}^T\mathbf{D}_v^T\mathbf{D}_v\mathbf{n})\end{bmatrix}.\end{aligned} \tag{10}$$

Two diagonal components have the same statistical properties. Therefore, we specifically examine the upper-left component. Letting $\xi(\mathbf{n}) = \mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_h\mathbf{n}$, we approximate the distribution of $\xi(\mathbf{n})$ by the gamma distribution to simplify the problem. Because the moment generating function (MGF) uniquely determine the distribution, we show the MGF of the variable $\xi(\mathbf{n})$ and the gamma distribution. The MGF of the variable $\xi(\mathbf{n})$ can be derived as

$$\begin{aligned}M_\xi(t) &= E(e^{t\xi(\mathbf{n})})\\ &= \int e^{t\cdot\mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_h\mathbf{n}}p_n(\mathbf{n})d\mathbf{n}\\ &= \int e^{t\cdot\mathbf{n}^T\mathbf{D}_h^T\mathbf{D}_h\mathbf{n}}\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}e^{-\frac{\mathbf{n}^T\mathbf{n}}{2\sigma^2}}d\mathbf{n}\\ &= \prod_{i=1}^{N}\frac{1}{(1-2\sigma^2 t\lambda_i)^{\frac{1}{2}}}\,,\end{aligned} \tag{11}$$

where λ_i is the i-th eigenvalue of the matrix $\mathbf{D}_h^T\mathbf{D}_h$. The MGF of the gamma distribution with the shape parameter α and scale parameter β is written as

$$M_g(t) = (\frac{1}{1-\beta t})^\alpha = \prod_{i=1}^{N}\frac{1}{(1-\beta t)^{\frac{\alpha}{N}}}\,. \tag{12}$$

Comparing Eq. (11) and Eq. (12), we approximate the MGF of the variable $\xi(\mathbf{n})$ by that of the gamma distribution with parameters:

$$\alpha = \frac{N}{2}\,, \quad \beta = \frac{2}{N}\sigma_n^2\mathrm{tr}(\mathbf{D}_h^T\mathbf{D}_h)\,, \tag{13}$$

where $\mathrm{tr}(\mathbf{D}_h^T\mathbf{D}_h)$ is the trace of matrix $\mathbf{D}_h^T\mathbf{D}_h$.

To select the weak textured patches, we define the null hypothesis as "the given patch is a flat patch with the white Gaussian noise" . We select the patches for which the null hypothesis is accepted. The null hypothesis is accepted if the maximum eigenvalue of the gradient covariance matrix is less than some threshold. The threshold τ is given with the given significance level δ and noise level σ_n as

$$\tau = \sigma_n^2 F^{-1}(\delta, \frac{N}{2}, \frac{2}{N}\mathrm{tr}(\mathbf{D}_h^T\mathbf{D}_h))\,, \tag{14}$$

where $F^{-1}(\delta, \alpha, \beta)$ is the inverse gamma cumulative distribution function with the shape parameter α and scale parameter β.

In the proposed weak textured patch selection algorithm, we select the patches of which the maximum eigenvalue of the gradient

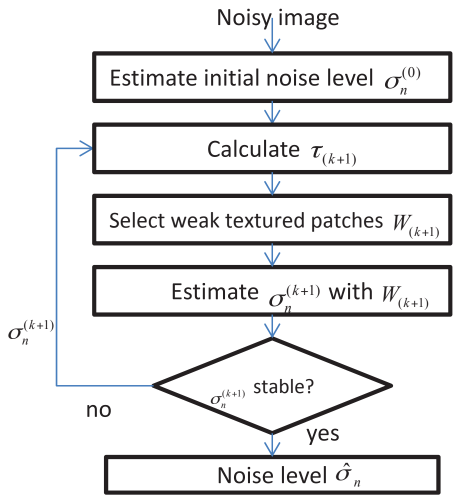


Fig. 2. Flowchart of the proposed iterative noise level estimation algorithm.

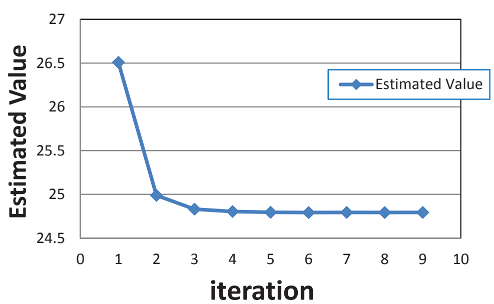


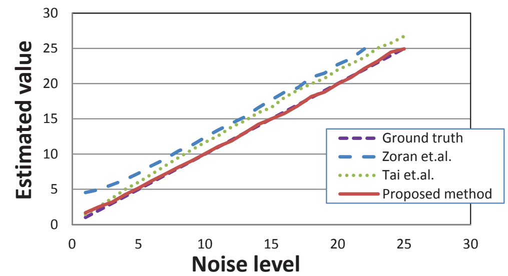
Fig. 3. Estimated noise level in each iteration (*mountain* image, true $\sigma_n = 25$)

covariance matrix is less than the threshold given in Eq.(14). The significant level δ must be given manually, for example 0.99. The noise level is also necessary to determine the threshold. In short, we can select the weak textured patches from the noisy image for the given noise level.

2.3. Iterative framework for noise level estimation

As discussed in Section 2.1, the noise level can be estimated easily if we can select the weak textured patches. To select the weak textured image patches, the noise level is required as described in Section 2.2. That poses a chicken-and-egg problem. To solve that chicken-and-egg problem, we introduce an iterative framework. Our iterative noise level estimation process is presented in Fig. 2.

First, an initial noise level $\sigma_n^{(0)}$ is estimated from the covariance matrix, which is generated using all patches in the input noisy image. Based on the k-th estimated noise level $\sigma_n^{(k)}$, the $(k+1)$-th threshold $\tau_{(k+1)}$ is determined. The weak textured patch dataset $W_{(k+1)}$ is selected from the noisy image using the threshold $\tau_{(k+1)}$. Then the $(k+1)$-th noise level $\sigma_n^{(k+1)}$ is estimated using selected $W_{(k+1)}$ with the threshold $\tau_{(k+1)}$. This process is iterated until the estimated noise level σ_n is unchanged. Although the convergence of this iteration process is not theoretically guaranteed, we have confirmed experimentally that this iteration process converges after several iterations. One example of the estimated noise level of each iteration is shown in Fig. 3.



(a) Noise level estimation result

(b) Original image (c) Weak textured part ($\sigma_n = 1$) (d) Weak textured part ($\sigma_n = 25$)

Fig. 4. Noise level estimation results of *mountain* image, where the yellow region represents selected weak textured patches.

3. EXPERIMENTAL RESULTS

We compare the proposed method [1] with results obtained using existing methods by different scenes with different noise levels. The estimation results are conducted directly from the noisy images for each noise level. The patch size we used is 7×7 and the noise level σ_n is from 1 to 25. One hundred natural images in the test set of Berkeley Segmentation Database (BSD) [8] are used for the experiments.

Fig. 4 portrays a scene of *mountain*, which contains both weak and rich textured regions. Because of the large amount of rich textures, the methods by Zoran and Weiss in [5] and Tai *et al.* in [3] overestimate the noise level. The proposed method can select and use only the weak textured region (yellow parts in (c) and (d)), the estimation results are more accurate as shown in (a). Fig.5 portrays a difficult scene called *gravel*. The whole image includes fine detail, which causes most methods to overestimate the noise level greatly. Since this image includes no obvious weak textures, in the low noise level cases, few patches are selected as shown in (c). As the noise level increases, e.g. $\sigma_n = 25$, even human vision is unable to distinguish the textures from the noise. Even though selected patches might contain some rich textured regions as (d) in this case, the proposed method achieves a better noise level estimation.

Table 1 shows the average, standard deviation, and root mean square error (RMSE) of estimated noise levels from the 100 images. The standard deviation reflects the ability of the estimator to deal with a variety of natural scenes. The RMSE is a good measure of precision for the estimator. From that comparison, significant improvement in the standard deviation and RMSE is apparent, which indicates that the proposed method is more accurate, more stable, and more scene-independent.

Noise level is an extremely important parameter for many image-processing applications. A typical one is blind denoising. BM3D, proposed by Dabov *et al.* [9], is one of recent state-of-the-art non-blind denoising algorithms. To verify the estimated noise

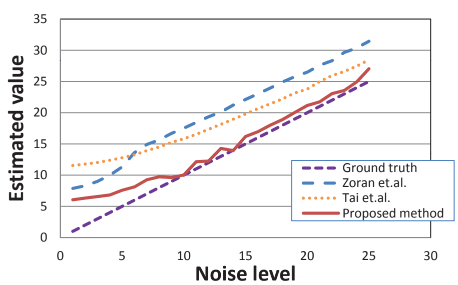
[1] MATLAB code is available on the author's webpage. `http://www.ok.ctrl.titech.ac.jp/res/NLE/WTP.html`

Table 1. Results for the Berkeley Segmentation Dataset (100 images), showing the average and the standard deviation of estimated noise levels, and root mean square error (RMSE) between the estimated noise level and the true noise level. Bold font represents better results.

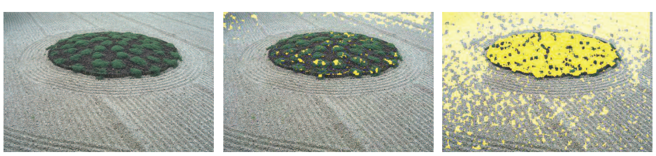
True noise level	Zoran and Weiss [5]			Tai *et al.* [3]			Proposed method		
	Average	Std. dev.	RMSE	Average	Std. dev.	RMSE	Average	Std. dev.	RMSE
1	2.129	1.662	1.986	2.059	1.574	1.890	**1.512**	**0.986**	**1.105**
5	**4.993**	1.464	1.455	5.761	1.083	1.320	5.285	**0.429**	**0.507**
10	**9.741**	1.589	1.602	10.635	0.825	1.038	10.282	**0.389**	**0.470**
15	14.643	1.606	1.637	15.553	0.677	0.871	**15.254**	**0.336**	**0.395**
20	19.582	1.645	1.689	20.482	0.613	0.778	**20.127**	**0.308**	**0.327**
25	24.464	1.649	1.725	25.458	0.549	0.713	**25.057**	**0.384**	**0.386**

Table 2. Average PSNR value of denoised image from the Berkeley Segmentation Dataset (100 images). The denoising algorithm is BM3D [9]. Bold font represents better results from estimated noise levels .

True noise level (PSNR)	Ground truth+ BM3D[9]		Zoran [5] + BM3D[9]		Tai [3] + BM3D[9]		Proposed method + BM3D[9]	
	Average	Std. dev.	Average	Std. dev.	Average	Std. dev.	Average	Std. dev.
1 (48.13)	49.094	0.703	44.141	3.588	47.424	3.050	**48.651**	**2.012**
5 (34.15)	37.432	1.743	36.966	2.171	36.830	2.155	**37.382**	**1.783**
10 (28.13)	33.133	2.151	32.965	2.305	32.824	2.392	**33.133**	**2.151**
15 (24.61)	30.880	2.340	30.789	2.430	30.719	2.496	**30.900**	**2.341**
20 (22.11)	29.417	2.457	29.363	2.520	29.326	2.562	**29.446**	**2.444**
25 (20.17)	28.368	2.520	28.340	2.553	28.368	2.548	**28.394**	**2.509**



(a) Noise level estimation results



(b) Original image (c) Weak textured part ($\sigma_n = 1$) (d) Weak textured part ($\sigma_n = 25$)

Fig. 5. Noise level estimation results of *gravel* image, where the yellow region represents selected weak textured patches.

level, here we apply the different estimated noise levels to the BM3D algorithm and obtain the PSNR value of the denoised image. Comparison of the results is shown in Table 2. The proposed method achieves the best result among the practically estimated noise levels.

4. CONCLUSION

As described in this paper, we proposed an algorithm to select weak textured patches from the images corrupted by the Gaussian noise. We applied the PCA technique to estimate the noise level based on the weak textured patch dataset. We use the maximum eigenvalue of the image gradient covariance matrix as the metric for texture strength and discuss how it changes with different noise levels σ_n. In contrast to state-of-the-art methods, the proposed method is more scene-independent and presents significant improvement for both accuracy and stability for a range of noise levels in various scenes.

5. REFERENCES

[1] K.Rank, M.Lendl, and R.Unbehauen, "Estimation of image noise variance," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 146, pp. 80, 1999.

[2] Aishy Amer, Amar Mitiche, and Eric Dubois, "Reliable and fast structure-oriented video noise estimation," *International Conference on Image Processing*, vol. 51, pp. I–840–I–843, 2002.

[3] Tai Shen-chuan and Yang Shih-ming, "A fast method for image noise estimation using laplacian operator and adaptive edge detection," *2008 3rd International Symposium on Communications, Control and Signal Processing*, pp. 1077–1081, March 2008.

[4] Shin Dong-hyuk, Park Rae-hong, Yang Seung-joon, and Jung Jae-han, "Block-based noise estimation using adaptive gaussian filtering," *IEEE Transactions on Consumer Electronics*, vol. 51, pp. 218–226, Feb. 2005.

[5] Zoran Daniel and Weiss Yair, "Scale invariance and noise in natural images," *2009 IEEE 12th International Conference on Computer Vision*, pp. 2209–2216, 2009.

[6] Christopher M. Bishop, "Pattern recognition and machine learning," *Springer*, p. 561, 2006.

[7] Zhu Xiang and Milanfar Peyman, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE trans. on image processing*, vol. 19, pp. 3116–32, 2010.

[8] D. Martin, D.Tal C.Fowlkes, and Jalik, "A database of human segmented natrual images and its application to evalueating segmentation algorithms and measuring ecological statistics," *8th Int'l Conf. Computer Vision*, vol. 2, pp. 416–423, 2001.

[9] Dabov Kostadin, Foi Alessandro, Katkovnik Vladimir, and Egiazarian Karen, "Image denoising by sparse 3-d transform-domain collaborative filtering.," *IEEE transactions on image processing*, pp. 2080–95, 2007.